A Calibration and registration of motion stage to transducer

A.1 Preliminaries

The optical tracking system measures position and rotational attitude of tracked rigid bodies in its field of view. Measurement of each tracked body returns a three-vector $\mathbf{v}$ and a 3x3 rotation matrix $\mathbf{R}$ (or a quaternion that can be converted to a rotation matrix). These define a transform between points in the local coordinate axes of the tracked body and the "absolute" defined by the camera tracking system or reference tracker:

$$\mathbf{x}_{abs} = \mathbf{R}\mathbf{x}_{loc} + \mathbf{v}$$

Unless otherwise stated, we will express the above rotation and translation operation as matrix $\mathbf{Q}$ using the longform 4x4 standard :

$$\mathbf{Q} = \begin{bmatrix} R_{00} & R_{01} & R_{02} & v_0 \\ R_{10} & R_{11} & R_{12} & v_1 \\ R_{20} & R_{21} & R_{22} & v_2 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

and vectors will be length-4 with a 1 appended as the last component: $\mathbf{x} = (x_0, x_1, x_2, 1)$. Using this convention a point is converted from local tracker coordinates to absolute coordinates by

$$\mathbf{x}_{abs} = \mathbf{Q}\mathbf{x}_{loc}$$

and from absolute to local by taking the inverse

$$\mathbf{x}_{loc} = \mathbf{Q}^{-1}\mathbf{x}_{abs}$$

A.2 Tracked hydrophone tip

A rigid body tracker (R) is mounted to a rod that holds the hydrophone, and a pivot calibration performed to dermine the tip offset. Let $\mathbf{t}_R$ denote the tip of the hydrophone returned from pivot calibration. After mounting the hydophone rod to the motion stage, a second tracker T is mounted to the transducer cone and optical tracking initiated. Acquisition from the tracking system returns the instantaneous pose of each tracker, $\mathbf{Q}_T$ and $\mathbf{Q}_R$. The tip position in absolute coordinates is $\mathbf{t}_{abs} = \mathbf{Q}_R\mathbf{t}_R$ and thus the position with respect to T is

$$\mathbf{t}_T = \mathbf{Q}_T^{-1}\mathbf{Q}_R\mathbf{t}_R$$

We can move the hydrophone to any position, record $\mathbf{Q}_T$ and $\mathbf{Q}_R$, and use this formula to localize the tip relatiave to calibration tracker T. Note: if tracking software permits T being used as a global reference during acquisition, then $\mathbf{Q}_T^{-1}$ is implicitly applied to the returned matrix $\mathbf{Q}_R$.

A.3 Motor-to-transducer registration

The motion stage was used to translate the hydrophone over a 3D grid of fiducial points to establish a transform between motor coordinates M and transducer coordinates T. The set of points (64 total) were first defined in the motor coordinates space. Let $F = \{\mathbf{x}_M^{(1)}, \mathbf{x}_M^{(2)}, \ldots \mathbf{x}_M^{(64)}\}$ be the set of fiducials in motor coordinates. At each point a series of $\mathbf{Q}_R$ and $\mathbf{Q}_T$ were recorded and averaged. We can now compute the corresponding hydrophone tip postion for each fiducial (n):

$$\mathbf{t}_T^{(n)} = (\mathbf{Q}_T^{-1})^{(n)} \mathbf{Q}_R^{(n)} \mathbf{t}_R$$

Doing this for all n gives a set of fiducial points in each sysem, $\{\mathbf{x}_M^{(n)}\}$ and $\{\mathbf{t}_T^{(n)}\}$. We solve for a rigid transform from M to T that minimizes error between the fidicual sets:

$$\mathbf{T}_{M,T} = \text{argmin}_{\mathbf{T}} \sum_{n=1}^{64} ||\mathbf{T}\mathbf{x}_M^{(n)} - \mathbf{t}_T^{(n)}||^2$$

such that

$$\mathbf{t}_T^{(n)} \approx \mathbf{T}_{M,T} \mathbf{x}_M^{(n)}$$

within fiducial registration error. Given an arbitrary point defined in motion stage coordinates, transform $\mathbf{T}_{M,T}$ estimates the attached hydrophone tip location in transducer coordinates.

Collection of fiducials is the crucial component in this calibration method. Spatial tracking errors or discrepancy between the real and brass hydrophones would show up as errors in the components of $\mathbf{T}_{M,T}$. An average $\mathbf{T}_{M,T}$ was collected in each calibration session by repeating the above procedure at a variety of camera positions encountered in live targeting. In cases where bias corrections were applied, $\mathbf{T}_{M,T}$ was modified by adding an offset $\mathbf{b}$ to the last column.

A.4 Beam voxels and affine

Beam mapping yields a 3D array of pressure values which we will call $P(i, j, k)$. The subscripts (i, j, k) form an index vector that refers to the address/position of a voxel in the array. Voxel indices can be converted to positions in the motion stage coordinate system (M) using an **affine transform**. Let matrix $\mathbf{M}$ be the motor affine, $\Delta x, \Delta y, \Delta z$ be the step sizes in each motor axis (0.4 mm) used during beam mapping, and x_0, y_0, z_0 be the position of the origin voxel $(i, j, k) = (0, 0, 0)$ in each motor axis. The motor affine is then

$$\mathbf{M} = \begin{bmatrix} \Delta x & 0 & 0 & x_0 \\ 0 & \Delta y & 0 & y_0 \\ 0 & 0 & \Delta z & z_0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

and a voxel (i, j, k) in the beam has motor position:

$$\mathbf{x}_M = \mathbf{M} \begin{bmatrix} i \\ j \\ k \\ 1 \end{bmatrix}$$

From the previous section we can transform this vector from motor coordinates into co-registered transducer coordinates

$$\mathbf{x}_T = \mathbf{T}_{M,T}\mathbf{x}_M = \mathbf{T}_{M,T}\mathbf{M} \begin{bmatrix} i \\ j \\ k \\ 1 \end{bmatrix} = \mathbf{A} \begin{bmatrix} i \\ j \\ k \\ 1 \end{bmatrix}$$

We can now store the calibration as an array of acoustic voxels $P(i, j, k)$ and an affine $\mathbf{A} = \mathbf{T}_{M,T}\mathbf{M}$. The NIFTI file-format was used to store voxels and affine, and is compatible with 3D Slicer.

A.5 Projection of the beam

In image-guided targeting, donut-shaped image fiducials are first used to determine the transform between pre-operative MR image coordinates (S) and tracked coordinates. Tracker (R) is the global reference during targing. Assume we have localized fiducials in each space, and let the physical-to-image transform be $\mathbf{T}_{R,S}$. The instantaneous transducer position and orientation during targeting is given by $\mathbf{Q}_T$. Then, for a point in transducer coordinates, the corresponding point in tracked coordinates is:

$$\mathbf{x}_R = \mathbf{Q}_T\mathbf{x}_T$$

(since R is the global reference, $\mathbf{Q}_R^{-1}$ has been implicitly applied to tracking output). Using the physical-to-image registration this position in the pre-operative image is

$$\mathbf{x}_S = \mathbf{T}_{R,S}\mathbf{x}_R = \mathbf{T}_{R,S}\mathbf{Q}_T\mathbf{x}_T$$

Combining this with the previous definition of acoustic affine, we have the full transformation to project the location of acoustic voxel $\mathbf{v} = (i, j, k, 1)$ into the image:

$$\mathbf{x}_T = \mathbf{T}_{M,T}\mathbf{M}\mathbf{v} = \mathbf{A}\mathbf{v}$$

and

$$\mathbf{x}_S = (\mathbf{T}_{R,S}\mathbf{Q}_T\mathbf{A})\mathbf{v}$$

$$\mathbf{v} = (\mathbf{T}_{R,S}\mathbf{Q}_T\mathbf{A})^{-1}\mathbf{x}_S$$

Using this transform acoustic voxels can be resliced into any arbitrary MR volume. For example in figures 5 and 8, beams were projected using tri-linear interpolation via the SciPy function `scipy.interpolate.interpn()`. First the locations of voxels in the target MR volume were compiled, $\{\mathbf{x}_S\}$, and transformed to a set of (real-valued/floating point) acoustic voxel indices $\{\mathbf{u}\}$ using the above equation. The set of index vectors corresponding to the beam map grid $P(i, j, k)$ was constructed and interpolated at the $\{\mathbf{u}\}$ points to yeild the re-sliced beam map projected into an MR volume.